Relevant database and user information

User ID	user_professor
User Password	qf5214_group4
Public IP	35.223.229.79
Connection(read&write)	supple-folder-418707:us-central1:qf5214project
Connection(read)	supple-folder-418707:us-central1:qf5214project-replica
Database name	Bitcoin_data
Database type	MySQL 8.0.31
Key file	qf5214_professor.json

First step of connection: Key file

Put the key file to load address like : C:\Users\User\Desktop\read_me\qf5214_professor.json

Second step of connection: Configuring the link engine

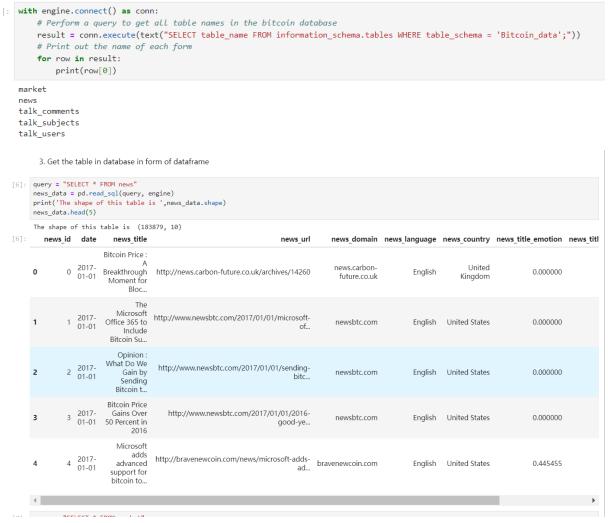Option 1 : jupyter notebook: 'connect_sql.ipynb' (you can get it in the fille '"read me")

1. Establish the connection with google cloud sql

```
[1]: #!pip install cloud-sql-python-connector
     # Firstly, installing this package to connect cloud sql!!!

[2]: from google.cloud.sql.connector import Connector, IPTypes
     import sqlalchemy
     from google.oauth2 import service_account
     from sqlalchemy import text
     import pandas as pd

[3]: # Specify the service account file path
     SERVICE_ACCOUNT_FILE = r'C:\Users\User\Desktop\read_me\qf5214_professor.json'

     # Create service account credentials
     credentials = service_account.Credentials.from_service_account_file(
         SERVICE_ACCOUNT_FILE
     )
     connector = Connector(credentials=credentials)

[4]: def getconn() -> sqlalchemy.engine.base.Connection:
         instance_connection_string = "supple-folder-418707:us-central1:qf5214project-replica"
         #instance_connection_string = "supple-folder-418707:us-central1:qf5214project"
         db_user = "user_professor"
         db_pass = "qf5214_group4"
         db_name = "Bitcoin_data"                          # Replace with your actual database name, but none seems to work

         conn = connector.connect(
             instance_connection_string,
             "pymysql",
             user=db_user,
             password=db_pass,
             db=db_name,
             ip_type=IPTypes.PUBLIC,  # Use public IP for access
         )
         return conn

     # Create the SQLAlchemy engine
     engine = sqlalchemy.create_engine(
         "mysql+pymysql://",  # Our database is MySQL8.0
         creator=getconn,
     )
```

Option 2 : spyder or pycharm: 'connect_sql.py' (you can also get it in the fille '"read me")

Third step of connection: read the database

Notes: please use 'Connection(read&write)' if want to change the data.

2. Get the name of table in database

```
with engine.connect() as conn:
    # Perform a query to get all table names in the bitcoin database
    result = conn.execute(text("SELECT table_name FROM information_schema.tables WHERE table_schema = 'Bitcoin_data';"))
    # Print out the name of each form
    for row in result:
        print(row[0])
```

```
market
news
talk_comments
talk_subjects
talk_users
```

3. Get the table in database in form of dataframe

```
query = "SELECT * FROM news"
news_data = pd.read_sql(query, engine)
print('The shape of this table is ',news_data.shape)
news_data.head(5)
```

The shape of this table is (183879, 10)

	news_id	date	news_title	news_url	news_domain	news_language	news_country	news_title_emotion	news_titl
0	0	2017-01-01	Bitcoin Price : A Breakthrough Moment for Bloc...	http://news.carbon-future.co.uk/archives/14260	news.carbon-future.co.uk	English	United Kingdom	0.000000	
1	1	2017-01-01	The Microsoft Office 365 to Include Bitcoin Su...	http://www.newsbtc.com/2017/01/01/microsoft-of...	newsbtc.com	English	United States	0.000000	
2	2	2017-01-01	Opinion : What Do We Gain by Sending Bitcoin t...	http://www.newsbtc.com/2017/01/01/sending-bitc...	newsbtc.com	English	United States	0.000000	
3	3	2017-01-01	Bitcoin Price Gains Over 50 Percent in 2016	http://www.newsbtc.com/2017/01/01/2016-good-ye...	newsbtc.com	English	United States	0.000000	
4	4	2017-01-01	Microsoft adds advanced support for bitcoin to...	http://bravenewcoin.com/news/microsoft-adds-ad...	bravenewcoin.com	English	United States	0.445455	

I would like to add that some other points:

'talk_comments' is a large table, and it takes a long time to export. On average, it takes about 5-15 minutes. You can choose to read a small amount to improve the speed, for example, only 100 rows are read.

Cloud sql may stop because the money in the account runs out, I can ensure that the cloud is available until April 30th, and the longest estimate can last until May 5th. We will not be able to access it after that, please forgive me.

Any connection question, please feel free to contact me by email e1124162@u.nus.edu, and I will try my best to answer your questions.